

Annex C. Methodological notes

Data harmonisation

Comparability of results across different countries is achieved by reproducing the same cleaning procedures and estimation techniques as used on the raw data. The statistical package produced by the OECD includes the following data cleaning instructions, applied to all country source data:

- Firms in sectors (1-digit NACE 2⁶) A, B, O, T, U are dropped.
- All negative values of variables that cannot plausibly be negative (employment, turnover, export revenues, labour costs, fixed costs, operating costs, material costs, depreciation, interest costs, research and development (R&D) expenditures, intangible fixed assets, tangible fixed assets, fixed assets, current assets, total assets, loans) are considered missing.
- Observations with rare values of the variable *year* are dropped (e.g. observations that refer to the future).
- The birth year value is considered missing if it is a later year than the latest year covered by the dataset.
- The accounting indicators based on ratios (e.g. value-added over turnover, intangible assets over tangible assets) are winsorized at 1% level (i.e. values below the first and above the 99th percentile of the distribution are replaced with the value of the 2nd and 98th percentile respectively).

The sector groups

Firms are classified within the industry-standard classification system (NACE 2 rev.2 sector classification). The regression analyses rely on the 4-digit classification. Sectoral analysis aggregates the NACE sections into six groups depending on the intensity of high-technology and knowledge (Table A C.1).

Table A C.1. Sectoral groups with corresponding NACE sector divisions

Aggregated sector groups	NACE Rev. 2 Divisions	Sector names
Low-tech and medium-low-technology manufacturing	10-12	Manufacture of food products, beverages and tobacco products
	13-15	Manufacture of textiles, apparel, leather and related products
	16-18	Manufacture of wood and paper products, and printing
	19	Manufacture of coke and refined petroleum products
	22, 23	Manufacture of rubber and plastics products, and other non-metallic mineral products
	24, 25	Manufacture of basic metals and fabricated metal products, except machinery and equipment
	31-33	Other manufacturing, and repair and installation of machinery and equipment
Medium-high and high-technology manufacturing	20	Manufacture of chemicals and chemical products
	21	Manufacture of pharmaceuticals, medicinal chemicals and botanical products
	26	Manufacture of computer, electronic and optical products
	27	Manufacture of electrical equipment

Aggregated sector groups	NACE Rev. 2 Divisions	Sector names
Knowledge-intensive services (KIS)	28	Manufacture of machinery and equipment
	29, 30	Manufacture of transport equipment
	50-51	Water transport; Air transport
	58-63	Information and communication
	64-66	Financial and insurance activities
	69-75	Professional, scientific and technical activities
	78, 80	Employment activities; Security and investigation activities
Less knowledge-intensive services	45-47	Wholesale and retail trade and repair of motor vehicles and motorcycles
	49	Land transport and transport via pipelines
	52-53	Warehousing and support activities for transportation; Postal and courier activities
	55-56	Accommodation and food service activities
	68	Real estate activities
	77	Rental and leasing activities
	79	Travel agency, tour operator, reservation service and related activities
	81-82	Services to buildings and landscape activities; Office administrative, office support and other business support activities
Education, social care and health services	85-88	Education; Human health and social work activities
Construction	41-43	Construction

Note: The manufacturing sectors are aggregated using Eurostat's high-technology classification of manufacturing industries (https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:High-tech_classification_of_manufacturing_industries). The service sectors are aggregated using Eurostat's definition of knowledge-intensive and less knowledge-intensive services.

Source: NACE division codes are extracted from https://ec.europa.eu/competition/mergers/cases/index/nace_all.html.

Identification of innovative activities in firms

The job functions that are used to approximate innovative activity in firms are classified depending on the keywords in the job function description. Types of classifications can vary across countries. For Finland and Portugal, job functions are classified using the International Standard Classification of Occupations (ISCO-08) classification.

The ISCO-08 classification provides a description of job functions with four different levels of detail for each job: major groups, sub-major groups, minor groups and unit groups. Each level includes a list of tasks typically required for the job.

The jobs that fall under the group “HR job functions” need to mention one or more of the following keywords: “human resource”, “career”, “training” and “staff development”.

These keywords correspond to ISCO occupation unit groups as follows:

- 1212 Human Resource Managers.
- 2423 Personnel and Careers Professionals.
- 2424 Training and Staff Development Professionals.
- 4416 Personnel Clerks.

Management can be tracked by identification of job function groups, which have “management”, “organisation” or “planning” in their title.

Management is represented by the following ISCO occupations:

- 1213 Policy and Planning Managers.
- 2421 Management and Organisation Analyst.
- 2422 Management Policy Specialist.

The digital technology job functions are all jobs that mention “information technology”, “multimedia”, “software”, “programmers”, “database”, “network” and “system” in their title or job functions.

These generate the following list of ISCO occupations:

- 1330 Information and Communications Technology Services Managers.
- 2356 Information Technology Trainers.
- Sub-major group 25:
 - 251 Software and Applications Developers and Analysts (2511 System Analysts, 5512 Software Developers, 2513 Web and Multimedia Developers, 2514 Applications Programmers, 2519 Software and Application Developers and Analysts).
 - 252 Database and Network Professionals (2521 Database Designers and Administrators, 2522 Systems Administrators, 2523 Computer Network Professionals, 2529 Database and Network Professionals).
 - Minor group 351 Information and Communications Technology Operations and User Support Technicians (3511 Information and Communications Technology Operations Technicians, 3512 Information and Communications Technology User Support Technicians, 3513 Computer Network and Systems Technicians, 3514 Web Technicians).

Research job functions are the jobs that mention “research” in the title of their profession (excluding 4227 Survey and Market Research Interviewers) and as well as “research” among the first tasks in the description of the occupation.

These are the following job functions by ISCO-08 classification:

- 1223 Research and Development Managers.
- 21 Science and Engineering Professionals.
- 2310 University and Higher Education Teachers.
- 2351 Education Methods Specialists.
- 2631 Economists.
- 2632 Sociologists, Anthropologists and Related Professionals.
- 2633 Philosophers, Historians and Political Scientists.
- 2634 Psychologists.

Marketing job functions typically have “marketing”, “advertising” or “public relations” in their job title. This includes the following jobs:

- 1221 Sales and Marketing Managers.
- 1222 Advertising and Public Relations Managers.
- 2431 Advertising and Marketing Specialist.
- 2432 Public Relations Specialist.

Each employee in the administrative dataset is assigned a value of one if their job belongs to one of the five groups of job functions, zero otherwise. The employee dataset is then aggregated on the firm level, such that it contains the firm identifier, year (if applicable) and a number of contracts/persons that worked in each of the five innovative job groups. Such an aggregated employment dataset is linked to balance sheet data for scaling-up analysis.

For other types of classification, cross-validation of identified job categories needs to ensure that analysis considers only relevant employees. Labelling of the job functions involves a correct and sensible translation of the keywords used in the data. Besides identifying the desired groups of job functions, other jobs with the same keywords may appear in a selection without characterising the targeted job category. For example, “management” can appear in titles or descriptions of non-managerial job functions.

Classification of workers by skill and education level

Education

The educational attainment of employees is categorised according to the European Qualifications Framework (EQF) classification. Employees are grouped in categories according to their highest-achieved education level as follows: i) less than high school education; ii) high school diploma; iii) undergraduate degree; and iv) at least a graduate university degree. In the case of countries where the education classification is different, the closest possible classification to the four categories is applied.

Skills

Employees are classified by the skill content of their occupations. Following the International Standard Classification of Occupations (ISCO-08), occupations are classified as high-skilled, medium-skilled and low-skilled (ILO, 2012^[1]). Occupation refers to the kind of work performed in a job. The concept of occupation is defined as a “set of jobs whose main tasks and duties are characterised by a high degree of similarity”. Skill is defined as the ability to carry out the tasks and duties of a given job and is a function of the complexity and range of tasks and duties to be performed in an occupation. Skill level is measured operationally by considering one or more of:

- The nature of the work performed in an occupation in relation to the characteristic tasks and duties defined for each ISCO-08 skill level.
- The level of formal education defined in terms of the International Standard Classification of Education (ISCED) (UNESCO, 1997^[2]) required for competent performance of the tasks and duties involved.
- The amount of informal on-the-job training and/or previous experience in a related occupation required for competent performance of these tasks and duties.

Table A C.2. Classification of occupations based on skill requirements

Broad skill level	ISCO-08
Skill levels 3 and 4 (high)	1. Managers
	2. Professionals
	3. Technicians and associate professionals
Skill level 2 (medium)	4. Clerical support workers
	5. Service and sales workers
	6. Skilled agricultural, forestry and fishery workers

Broad skill level	ISCO-08
	7. Craft and related trades workers
	8. Plant and machine operators, and assemblers
Skill level 1 (low)	9. Elementary occupations
Armed forces	0. Armed forces occupations
Not elsewhere classified	X. Not elsewhere classified

Calculation of firm wage premium and the firm gender gap

The wage premium and gender wage gap of high-growth firms compared to non-high-growth firms are based on differences in residual wages. The estimation procedure follows three steps:

1. A Mincer wage regression is estimated, where log hourly wages are regressed on a set of worker characteristics, namely *tenure* and *tenure*², *age* and *age*², *education*, *occupation* and *sector*, using data from all workers in all firms. Based on the estimated coefficients from this regression, wage residuals for each individual are estimated by subtracting estimated from actual log wages. This provides estimates of the wage component of individual workers that is not explained by individual characteristics. This procedure controls for potential differences in the workforce composition with respect to individual characteristics across genders or scalars and non-scalars.
2. Two types of mean residual wages are computed for each individual firm. The first type assesses the average wage premium and is computed for all workers in a firm. The second type is computed separately for all male and female workers in a firm to assess the gender wage gap. The average gender wage gap in a firm is then computed by calculating the difference between the average residual earnings of male and female workers in a given firm.
3. Average residual wages and average residual gender wage gaps compare the wage differences between scalars and non-scalars.

References

- ILO (2012), "International Standard Classification of Occupations 2008 (ISCO-08): Structure, group definitions and correspondence tables", http://www.ilo.org/global/publications/ilo-bookstore/order-online/books/WCMS_172572/lang--en/index.htm (accessed on 9 June 2021). [1]
- UNESCO (1997), *International Standard Classification of Education - ISCED 1997*, http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-1997-en_0.pdf. [2]

Notes

¹ The probability is estimated with a linear probability model (LPM) at the firm level. The dependent variable is a binary indicator equal to one if the firm is in the final year of a three-year scaling-up period and equal to zero otherwise. The dependent variables are size class, age class and 2-digit sector categorical variables. The model also includes year fixed effects.

² Earnings before interest, taxes, depreciation.

³ CADS is a proprietary database administered by CERVED Group Ltd. for credit risk evaluation. It collects detailed balance sheet and income statement information of non-financial corporations since 1982 and it is the largest sample of Italian firms for which financial data are observed.

⁴ Neither INPS data nor the labour cost from CADS include independent and agency workers. Information on them cannot be retrieved from either of the two datasets. Although this is a strong limitation, integrating these workers would require estimating statistical imputation rules on a third dataset, and dataset of all active firms, access to which is currently restricted.

⁵ The SABI (Iberian Balance-Sheet Analysis System) is owned by the market research company Informa-Bureau van Dijk (<http://www.informa.es/en>) and constitutes the Spanish input to the Amadeus and Orbis datasets.

⁶ Nomenclature of Economic Activities in the European statistical classification of economic activities.



From:

Understanding Firm Growth

Helping SMEs Scale Up

Access the complete publication at:

<https://doi.org/10.1787/fc60b04c-en>

Please cite this chapter as:

OECD (2021), "Methodological notes", in *Understanding Firm Growth: Helping SMEs Scale Up*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/b42bbdff-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.